DOCUMENT RESUME

ED 237 174 JC 830 839

AUTHOR	Doyle, Kenneth O., Jr.
TITLE	Evaluating Teaching.
PUB DATE	30 Sep 83
NOTE	4p.; Excerpted from "Evaluating Teaching," Heath and Company.
PUB TYPE	Collected Works - Serials (022) -- Viewpoints (120) -- Guides - Non-Classroom Use (055)
JOURNAL CIT	Innovation Abstracts; v5 n27 Sep 30 1983
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Academic Freedom; *Accountability: Beliefs; *College Instruction; College Students; Evaluation Criteria; Faculty College Relationship; Instructional Improvement; Postsecondary Education; Student College Relationship; *Teacher Evaluation
IDENTIFIERS	*Evaluation Problems; Evaluation Utilization; PF Project

ABSTRACT

One philosophical issue underlying instructional evaluation arises from tension between the concepts of academic freedom and academic responsibility. Academic freedom can be cited as an argument against evaluating teaching in circumstances where evaluation genuinely encumbers the pursuit or dissemination of knowledge. Another source of tension centers around the notion that instructor's credentials, position, and expertise preclude evaluation. Some feel that until we know more about evaluation, we should do nothing. Others feel that failure to evaluate may lead to sins more grievous than would evaluation with uncertainty. In the face of increasing demands placed upon faculty to furnish more and more evidence supporting the amount, quality, and value of what they do, many faculty members want to do away with intrusive evaluation policies, while students and administrators push for objective data to facilitate decisions and guarantee educational quality. A major psychometric issue in instructional evaluation is whether evaluation should focus on instructional outcomes or on instructional processes. The problems of evaluating teaching by measuring student learning are both philosophical and methodological. On the other hand, evaluating teaching by measuring course and instructor characteristics and behavior that are indicative of good teaching is no less complex. A reasonable goal would be to devise an evaluation plan that takes advantage of the strengths of process measures, but that also guards against their weaknesses. (LH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
X Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Karen Watkins
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED237174

EVALUATING TEACHING

By
Kenneth O. Doyle, Jr.
University of Minnesota

September 30, 1983

From: Innovation Abstracts, Vol. V, No. 27
National Institute for Staff and Organizational Development
The University of Texas at Austin

JC 830 839

INNOVATION ABSTRACTS

VOL. V
NO. 27

Published by the National Institute for Staff and Organizational Development
With support from the W. K. Kellogg Foundation

EVALUATING TEACHING

Philosophical Issues

Perhaps the most important philosophical issue underlying instructional evaluation arises from a certain tension between the concepts of academic freedom and academic responsibility.

Academic freedom is sometimes cited as an argument against evaluating teaching, apparently on the grounds that evaluation often suggests change and change that is not entirely spontaneous is not free. However, a suggestion to change is not necessarily a requirement to change, and to extend the principle of academic freedom from the substance to the manner of teaching strains the traditional concept. Moreover, to use the principle of academic freedom as an unqualified reason for not evaluating teaching or, worse, as an excuse for poor teaching, risks doing violence to other people's rights under the principle of academic responsibility.

There may be circumstances, however, under which the academic freedom argument against evaluating teaching is valid: when evaluation genuinely encumbers the pursuit or dissemination of knowledge. Thus the argument could be valid if, say, anxiety about a peer review or a student survey were to inhibit an instructor from teaching some important course material. Although the inhibition may be more the result of the instructor's anxiety than of the evaluation itself, it is possible for evaluation to be used as a weapon against instructors who teach, for example, politically sensitive courses. In bona fide cases of this sort, the institution needs to assure that the academic freedom of neither teacher, nor students, nor colleagues is violated and that the academic responsibilities of all are fulfilled. This is perhaps best achieved in a deliberative evaluation process, in which the instructor, review committee members, and administrators can take into account any violations of either principle.

A second philosophical issue is, do the instructor's credentials, position, and expertise preclude evaluation? One occasionally encounters the argument that, especially in the sanctity of the classroom, there is no authority beyond the instructor, and hence no evaluation of the instructor that could be appropriate. For the most part, however, this argument fails to persuade. On the contrary one could argue with greater force that teaching must be evaluated, if only to confirm or disconfirm these claims of privilege.

A final philosophical issue reflects academe's inherent cautiousness and perhaps perfectionism or fear of error: Must instructional evaluation await a perfect evaluation technology? Teaching is still imperfectly understood and the instruments of evaluation are in some ways primitive. Faculty, then, who are often trained critics, sometimes direct their considerable critical powers at these deficiencies and propose that "until we know more, we should do nothing." But evaluation can proceed if teaching is sufficiently well understood and evaluation technology sufficiently well advanced that teaching can be responsibly albeit imperfectly assessed. Indeed, failure to evaluate may lead to sins more grievous than would evaluation with uncertainty. Instructional and personnel decisions are going to be made in any event, and these decisions are likely to be better the more they are enlightened by responsible evaluation procedures.

Psychodynamic Issues

Most psychodynamic issues in instructional evaluation reduce to the notion of control. Faculty value doing what they want to do in the way they want to do it,. In recent years, however, faculty have been pressed more and more to change their work styles, often by taking on tasks they find distasteful or even threatening and giving up pursuits they hold dear. Moreover, faculty are being required to furnish more and more evidence supporting the amount and quality, and even the value, of what they do. At the same time, they find rewards diminishing--salaries stagnate, grant sources disappear, time for reflection is scarce, appreciation of their effort is rarely expressed. Lurking beneath these external pressures are internal stresses as well. Faculty wonder if, through all of this evaluation, their worst and deepest fears will finally be realized: they will be found out.

Administrators and students are experiencing much the same turmoil. For administrators, every budget cut brings painful decisions; every move is criticized; every decision challenged, even in the courts. For students the cost of education continues to rise, the rewards are in serious doubt, and there is little they can do about it. In short, faculty, administrators, and students often feel they are losing control of their lives.

One effect of this feeling of loss of control is a drive to put things in order. For faculty, this might mean a pressing desire to return to a secure and simpler collegiality without the intrusiveness of evaluation policies and

Program in Community College Education, The University of Texas at Austin, EDB 348, Austin, Texas 78712

procedures and without such painful responsibilities as the rejection of junior colleagues seeking tenure. For administrators the drive for order might manifest itself in a thirst for objective data to facilitate decisions and remove the burden of subjectivity. For students it might mean a need for confidence in the quality of the education that is offered them and a passion for procedures and mechanisms to guarantee that quality.

The problem is that these drives compete, and efforts to put one group's life in order may further unsettle another's--and there is no ready solution to this conflict of needs. Refining evaluation instruments is desirable, but overall a simplistic response to deeply held anxieties. Clarifying policies and requirements is useful, but also misses the mark. Providing all the avenues for grievance in the world is still at best a superficial response. Perhaps the most one can do is try to recognize the other people's needs, show a little empathy, improve what can be improved, and press on.

Psychometric Issues

A major psychometric issue in instructional evaluation is, should the evaluation focus on instructional outcomes or on instructional processes. Instructional outcomes are the results of teaching, the diverse aspects of student learning. Process variables are the course and instructor characteristics and instructor behaviors that lead to student learning. The purpose of teaching is to facilitate learning, and the goal of instructional evaluation is to identify who or what leads to learning. The issue is whether one can approach this goal directly, by measuring learning, or must approach it indirectly, by measuring course and instructor characteristics.

The problems of evaluating teaching by measuring student learning are both philosophical and methodological. Philosophically, it seems unfair to evaluate a person in terms of what other people do or fail to do, at least at the beginning of the course, and student motivation throughout the course. If the test results were to be used to compare one instructor with another, the testing procedure would need to ensure that student differences in ability, motivation, and prior achievement were controlled for by random assignment of students to sections (which is sometimes inconvenient, usually impossible) or by application of statistical adjustments (which do not work so well in practice as they do in theory, see Glass, 1974). Second, the tests would need to be identical or equivalent in content and difficulty in all courses or sections, so that one instructor did not seem a less effective teacher than another simply because the latter gave an easier test. Third, there would need to be some way for instructors to assure themselves (and their colleagues) that the tests did indeed measure the depth and breadth of what their students were supposed to learn, so that people would not be misled into believing that a teacher is effective merely because the students successfully memorized a collection of facts. Finally, there would need to be some central mechanism to make sure that all the necessary controls were in place and to receive and act on appeals.

On the other hand, there is a massive literature that shows that good achievement tests can be derived, and faculty in general do seem to consider even routine classroom tests good enough for use in important decisions about students. Moreover, it is clear that instructors do have some responsibility for whatever learning occurs in their course. To be reasonable, then, the evaluation plan should take advantage of the strengths of classroom tests without losing sight of their weaknesses.

Evaluating teaching by measuring course and instructor characteristics and behaviors that are indicative of good teaching is certainly no less complex. Some seventy years of intense research effort have failed to produce a widely accepted, empirically supported theory or definition of good teaching. Recent research has made matters even more difficult by pointing out the importance of attending to the differential effects of teaching methods on different kinds of students. As unacceptable as it would be to reward or punish instructors on the basis of the dubious results of classroom tests, it would be equally unacceptable to reward or punish them for their conformity to a list of qualities that may have nothing to do with promoting learning in their courses.

But the negative case has again been overstated. Instructional research has made significant progress toward identifying the important features of good teaching, and certain evaluation questionnaire items have been shown to be valid. As is the case with classroom testing, a reasonable goal would be to devise an evaluation plan that takes advantage of the strengths of process measures, but that also guards against their weaknesses.

Kenneth O. Doyle, Jr.
Research Associate
Measurement Services Center
University of Minnesota

Excerpted from *Evaluating Teaching*. (Lexington, Massachusetts: Lexington Books, D. C. Heath and Company. c 1983, D. C. Heath and Company.)

Managing Editor, Karen Watkins
September 30, 1983, VOL. V, NO. 27

INNOVATION ABSTRACTS is a publication of the National Institute for Staff and Organizational Development, EDB 348, The University of Texas at Austin, Austin, Texas 78712, (512) 471-7545. Subscriptions are available to nonconsortium members for $35 per year. Funding in part by the W. K. Kellogg Foundation and the Fund for the Improvement of Post-Secondary Education. Issued weekly when classes are in session during fall and spring terms and bimonthly during summer months.

© The University of Texas at Austin, 1983
Further duplication is permitted only by MEMBER institutions for their own personnel